

DATA PROCESSOR

Patent Number: JP10275157
Publication date: 1998-10-13
Inventor(s): YODA NAOYUKI
Applicant(s):: SANYO ELECTRIC CO LTD
Requested Patent: JP10275157
Application Number: JP19970080534 19970331
Priority Number(s):
IPC Classification: G06F17/30.; G06F13/00 ; G06F17/28
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To reduce a physical, a mental, and an economical burden that retrieval of a great large amount of network information places on a user and to evade display or translation of unpleasant information on a malediction, a slander, etc., and further information on violent contents, obscene contents, etc., that do not desire to be shown to children.

SOLUTION: On a data processor where a program like a browser runs, a feature data DB 1 which stores feature data showing features of specific information, and a program including a feature data detecting means which detects whether or not feature data are included in retrieved information and a display control means which inhibits the above-mentioned feature data from being displayed by the retrieving and display means when the retrieved information includes the above-mentioned feature data are installed from, for example, a CD-ROM and run.

Data supplied from the esp@cenet database - I2

TOP

E6308

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-275157

(43)公開日 平成10年(1998)10月13日

(51) Int.Cl.⁶
 G 0 6 F 17/30
 13/00 3 5 4
 17/28

F I
 G 0 6 F 15/40 3 2 0 B
 13/00 3 5 4 D
 15/38 Z
 15/40 3 1 0 F

審査請求 未請求 請求項の数 3 O.L. (全 6 頁)

(21)出願番号 特願平9-80534

(71)出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(22)出願日 平成9年(1997)3月31日

(72)発明者 余田 直之

大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内

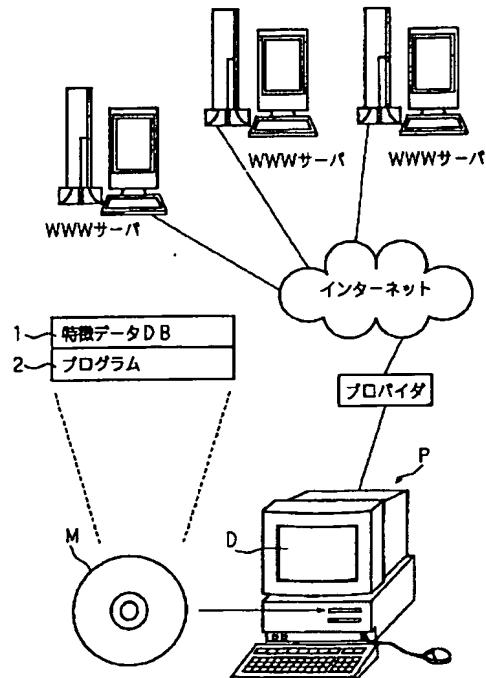
(74)代理人 弁理士 河野 登夫

(54)【発明の名称】 データ処理装置

(57)【要約】

【課題】 嶄大な量のネットワーク情報の検索がユーザに与える肉体的・精神的・経済的負担を軽減し、また、誹謗、中傷等の不愉快な情報、さらに、子供に見せたくない暴力的内容、猥褻な内容の情報等の表示又は翻訳を回避する。

【解決手段】 ブラウザのようなプログラムが動作するデータ処理装置に、所定の情報の特徴を表す特徴データを格納する特徴データDB1と、検索した情報に該特徴データが含まれているか否かを検出する特徴データ検出手段、及び検索した情報に前記特徴データが含まれている場合は、前記検索表示手段による該情報の表示を禁止する表示制御手段を含むプログラムとを、例えばCD-ROMからインストールして動作させる。



【特許請求の範囲】

【請求項1】 ネットワークが提供している情報を検索して表示する検索表示手段を備えたデータ処理装置において、所定の情報の特徴を表す特徴データを格納するデータベースと、検索した情報に該特徴データが含まれているか否かを検出する特徴データ検出手段と、検索した情報に前記特徴データが含まれている場合は、前記検索表示手段による該情報の表示を禁止する表示制御手段とを備えたことを特徴とするデータ処理装置。

【請求項2】 ネットワークから検索した情報を翻訳する翻訳手段を備えたデータ処理装置において、所定の情報の特徴を表す特徴データを格納するデータベースと、検索した情報に該特徴データが含まれているか否かを検出する特徴データ検出手段と、検索した情報に前記特徴データが含まれている場合は、前記翻訳手段による該情報の翻訳を禁止する翻訳制御手段とを備えたことを特徴とするデータ処理装置。

【請求項3】 前記データベースに特徴データを登録する手段を備えた請求項1又は2記載のデータ処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、例えばインターネットのWWW (World-Wide Web)が提供する情報を検索して表示し、また翻訳する機能を備えたデータ処理装置に関する。

【0002】

【従来の技術】最近のインターネットの人気によって、例えばWWWが提供している情報を検索するユーザが増加している。インターネットには国境がなく、また現状では法律が確立しておらず、誰でも容易に情報を発信することができるので、インターネットが提供する情報は膨大な量に上っている。このようなインターネットが提供する情報を検索して表示するソフトウェアとして、多数のブラウザが販売されている。

【0003】また、インターネットで流通している情報のほとんどは英語であるため、これを日本語へ翻訳するソフトウェアの要求が高く、多数の翻訳ソフトウェアが販売されている。

【0004】

【発明が解決しようとする課題】以上のように、インターネットが提供する量は膨大であるので、何らかの条件を設定して検索対象を絞り込んだとしても、検索される情報の量は大量になる。従って、検索した情報を表示することが、ユーザに大きな肉体的・精神的負担を与え、さらに検索した情報の表示に長時間を要するので、インターネットへの接続時間が長期化して経済的負担も与える。

【0005】また、インターネットへの情報の提供者は匿名が可能があるので、誹謗、中傷、また暴力的、猥褻、非合法な内容の情報も提供されている。従って、テキスト以外に画像、音声も提供しているインターネットから、暴力的、猥褻な内容の情報が家庭環境に侵入した場合、子供がそれらを目にする可能性が高い。

【0006】さらに、インターネットから提供される情報は、情報のジャンル、専門度、文体等が多種多様であるので、そのような情報に前編集のような加工を加えず翻訳した場合、高品質の翻訳が得られないのが現状である。

【0007】本発明はこのような問題点を解決するためになされたものであって、検索したネットワーク情報を表示する前、又は翻訳する前に、設定した条件に一致するような、表示したくない情報、内容を知りたくない情報をフィルタリングして表示対象、又は翻訳対象から除外することにより、ネットワーク情報の検索がユーザに与える肉体的・精神的・経済的負担を軽減し、また、誹謗、中傷等の不愉快な情報、さらに、子供に見せたくない暴力的内容、猥褻な内容の情報等の表示又は翻訳を回避し、ユーザのネットワーク環境を快適にするデータ処理装置の提供を目的とする。

【0008】

【課題を解決するための手段】第1発明のデータ処理装置は、ネットワークが提供している情報を検索して表示する検索表示手段を備えたデータ処理装置において、所定の情報の特徴を表す特徴データを格納するデータベースと、検索した情報に該特徴データが含まれているか否かを検出する特徴データ検出手段と、検索した情報に前記特徴データが含まれている場合は、前記検索表示手段による該情報の表示を禁止する表示制御手段とを備えたことを特徴とする。

【0009】第2発明のデータ処理装置は、ネットワークから検索した情報を翻訳する翻訳手段を備えたデータ処理装置において、所定の情報の特徴を表す特徴データを格納するデータベースと、検索した情報に該特徴データが含まれているか否かを検出する特徴データ検出手段と、検索した情報に前記特徴データが含まれている場合は、前記翻訳手段による該情報の翻訳を禁止する翻訳制御手段とを備えたことを特徴とする。

【0010】第3発明のデータ処理装置は、第1又は第2発明に加えて、前記データベースに特徴データを登録する手段を備えたことを特徴とする。

【0011】本発明のデータ処理装置は、誹謗、中傷等の不愉快な情報、暴力的、猥褻的、非合法な内容の情報を特徴付けているキーワード、そのような情報の発信者のURL (Uniform Resource Locator)のような、情報の特徴を表す特徴データをデータベース化しておき、ネットワークから検索した情報のうち、特徴データを含む情報を表示対象又は翻訳対象から除外する。

【0012】これによって、ネットワーク情報の検索がユーザーに与える肉体的・精神的・経済的負担を軽減し、また、誹謗、中傷等の不愉快な情報、さらに、子供に見せたくない暴力的内容、猥褻な内容の情報等の表示又は翻訳を回避することができる。

【0013】また本発明のデータ処理装置は、ユーザーは任意の特徴データをデータベースに登録する。これによって、多くのユーザーが共通して目にしたくない情報の特徴データ以外に、例えば翻訳不可能な専門用語が多くて訳語が得られず、低品質の翻訳しか得られないと予めわかっているようなジャンルの情報は翻訳しないといったように、ユーザー固有の特徴データを登録することでユーザーの仕様に合った検索環境の構築が可能になる。

【0014】

【発明の実施の形態】図1は本発明のデータ処理装置の構成を示す模式図である。本形態では、汎用のパーソナルコンピュータPのドライブに、後述する特徴データベース及びフィルタリング関連のプログラムが記録されているCD-ROMのような記録媒体Mが装填されて本発明のデータ処理装置として動作し、ディスプレイDに検索結果、翻訳結果等を表示するものとする。また、パーソナルコンピュータPはプロバイダを介してインターネットに接続されており、WWWサーバが提供する各種の情報画面を検索できるようになっている。なお、以下の説明ではインターネットで広く流通しているHTMLテキストをフィルタリングする場合を例とする。

【0015】記録媒体Mには、フィルタリングによって表示対象、翻訳対象から除外したい誹謗、中傷を含む情報、暴力的、猥褻、非合法な内容の情報等の特徴を表すキーワード、そのような情報の発信者のURL等が登録された特徴データDB(データベース)1が記録されている。なお、特徴データは装置側であらかじめ登録したものをユーザーに提供することも、いくつかの特徴データの候補の中からユーザーを選択させて登録することも、またユーザーが任意の特徴データを登録することも可能である。

【0016】また記録媒体Mにはフィルタリング関連のプログラムが記録されている。図2(a)の例では、WWWを検索して情報の内容を表示するブラウザ21、検索したHTMLテキストを解析して、特徴データDB1に登録されている特徴データがHTMLテキストから検出されるか否かを判定する特徴データ検出モジュール22、及び特徴データ検出モジュール22によって特徴データが検出された場合はブラウザ21にテキストを表示させないように制御する表示制御モジュール23が記録されている(図2(a))。

【0017】図2(b)の例では、上述のブラウザ21、特徴データ検出モジュール22に加え、辞書を参照して原文の形態素解析・構文解析・構文変換・形態素生成を実行して訳文を生成する翻訳モジュール24と、特徴データ検出モジュール22によって特徴データが検出された場合は

翻訳モジュール24にテキストを翻訳させないように制御する翻訳制御モジュール25とが記録されている。

【0018】図2(c)(d)の例では、図2(a)(b)に示すプログラムに加えて、例えば、後述のように、テキストの表示画面に設けた登録ボタンBが押された場合に、表示しているテキストにおいて出現頻度が高いキーワード、ハイパーアリンク先のURL等を特徴データDB1に登録する特徴データ登録モジュール26がそれぞれ記録されている。

【0019】次に、本発明のデータ処理装置の動作を、図3のフローチャート、図4の特徴データDBの概念図、図5のHTMLテキストの一例、及び図6のHTMLテキストの解析の概念図に基づいて説明する。図5に示すようなHTMLテキストをWWWから入力し(ステップS1)、その1文を入力する(ステップS2)。入力した1文を、図6に示すように、HTML情報(タグ)とテキスト情報とに分離するHTML解析を行う(ステップS3)。HTML解析した1文の文字列を意味を持つ最小の言語単位である形態素に分解し、分解した語の品詞、活用形等を明らかにする形態素解析を行う(ステップS4)。

【0020】次に、テキストへの出現頻度が所定回数以上であるキーワード、ハイパーアリンク先のURL等をテキストの特徴として抽出し(ステップS5)、以上のステップS2~S5を、WWWから入力したテキストの全文について繰り返す(ステップS6)。

【0021】抽出した特徴データを特徴データDB1の蓄積データと照合し(ステップS7)、抽出した特徴データが特徴データベースDBに登録されているか否かをチェックする(ステップS8)。抽出した特徴データが特徴データDB1に蓄積されていない場合はディスプレイDに表示し、又は翻訳処理を行う。一方、抽出した特徴データが特徴データDB1に蓄積されていた場合は、表示又は翻訳を行わず、処理を終了する。

【0022】例えば、図5に示すHTMLテキストを図6のようにタグとテキスト情報とに分離した結果、特徴データDB1に登録されている“nincompoop(まぬけ)”というキーワードが2回、“blah(ばかばかしいこと)”が3回テキスト情報に出現している。また、HTMLテキストでは、タグ「<A href」とタグ「>」とで挟んだファイルにハイバージャンプするが、そのURLが特徴データDB1に登録されているURLと一致する。従って、このテキストを表示対象又は翻訳対象から除外する。

【0023】次に、特徴データを登録する場合について、図7の表示画面の一例に基づいて説明する。WWWを検索した結果、図7のように表示された図5の内容を持つテキストをユーザーが見て、次回の検索からこのテキストを表示対象又は翻訳対象から除外したい場合、図7の表示画面において、例えばソフトボタンである登録ボタンBをユーザーが押すと、そのテキスト内で出現頻度が高いキーワード、そのテキストの発信者のURL、又はテキ

ストのハイパージャンプ先のURLが、図5に示すような状態で特徴データDB1に登録される。

【0024】なお、本例では、特徴データDB、及び各種プログラムが同一の記録媒体に記録されている形態について説明したが、これらは別個の記録媒体に記録されていてもよい。

【0025】また、特徴データDBだけが、又は特徴データDBと各種プログラムとの両方が、又は一部のモジュールが装置に予めインストールされている形態であっても同様の効果が得られる。

【0026】さらに、記録媒体MはCD-ROMに限るものではなく、他の記録媒体の使用も可能であって、さらにこれらのプログラムがオンラインで提供される形態であっても同様の効果が得られる。

【0027】

【発明の効果】以上のように、本発明のデータ処理装置は、検索したネットワーク情報を表示する前、又は翻訳する前に、設定した条件に一致するような、表示したくない情報、内容を知りたくない情報をフィルタリングして表示対象、又は翻訳対象から除外するので、ネットワーク情報の検索がユーザに与える肉体的・精神的・経済的負担を軽減し、また、誹謗、中傷等の不愉快な情報、さらに、子供に見せたくない暴力的内容、猥褻な内容の情報等の表示又は翻訳を回避し、ユーザのネットワーク環境を快適にするという優れた効果を奏する。

【図面の簡単な説明】

【図4】



キーワード	URL
nинкомпуп	here.htm
blah	hymn.htm

【図5】

```

<html>
<body>
<title>You are a nincompoop!</title>
<center><h1> You are a nincompoop!</h1></center>
<center><h2>Blah Blah Blah</h2></center>
<h2>You fell off the stairs yesterday.</h2>
Click<a href="www.acme.com/here.htm">HERE</a>for a really funny picture.
<b>You study Klingon literary works in translation: <a href="www.acme.com/hymn.htm">
Vedic hymns. </a></b><h3>
</body>
</html>

```

【図1】本発明のデータ処理装置の構成を示す模式図である。

【図2】本発明のデータ処理装置に格納されるプログラムの例を示した図である。

【図3】本発明のデータ処理装置の動作のフローチャートである。

【図4】本発明のデータ処理装置の特徴データDBの概念図である。

【図5】HTMLテキストの一例を示す図である。

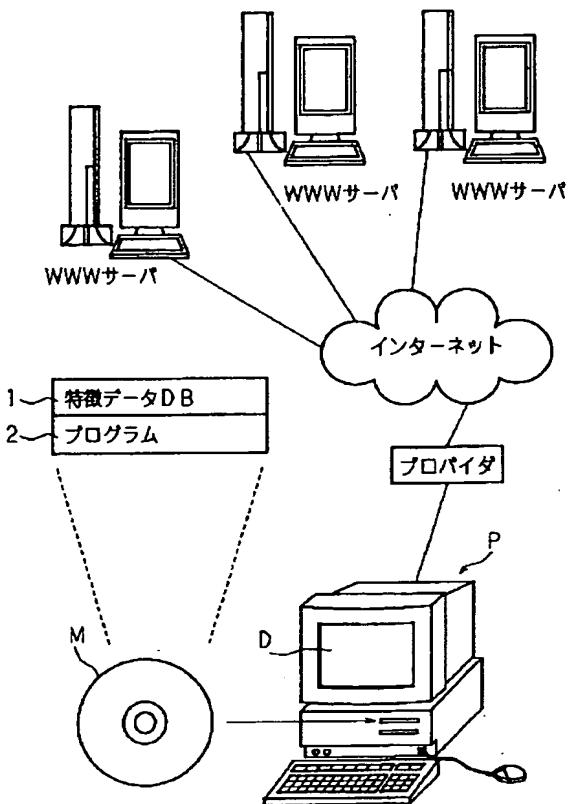
【図6】HTMLテキストの解析の概念図である。

【図7】本発明のデータ処理装置における検索結果の表示画面の一例の図である。

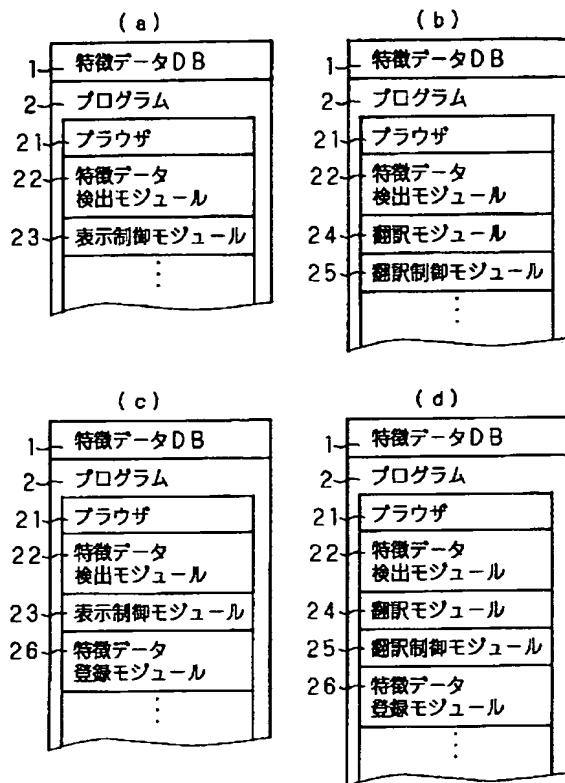
【符号の説明】

- 1 特徴データDB
- 2 プログラム
- 21 ブラウザ
- 22 特徴データ検出モジュール
- 23 表示制御モジュール
- 24 翻訳モジュール
- 25 翻訳制御モジュール
- 26 特徴データ登録モジュール
- B 登録ボタン
- D ディスプレイ
- M 記録媒体
- P パーソナルコンピュータ

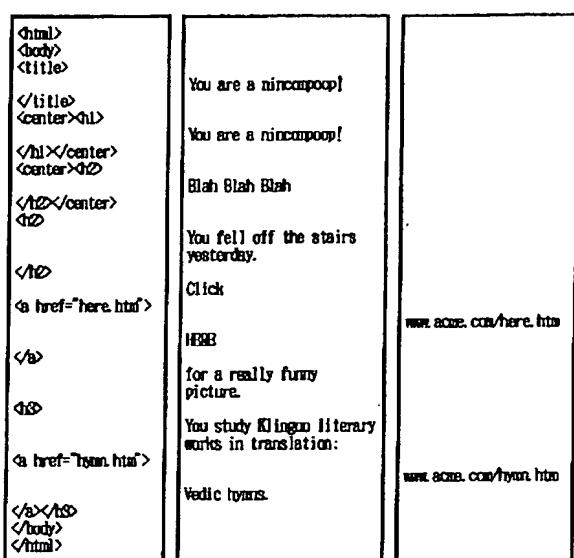
【図1】



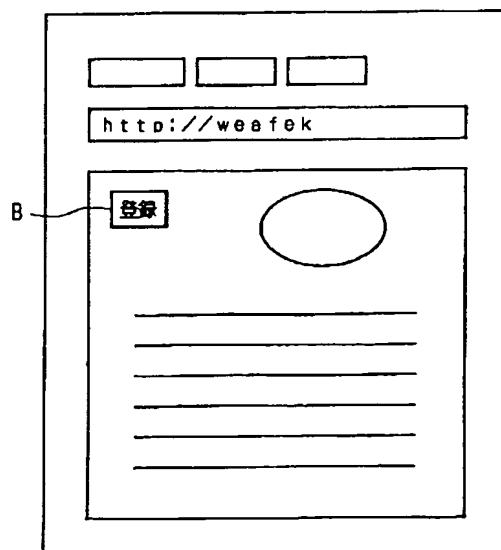
【図2】



【図6】



【図7】



【図3】

